

EXOPLANETARY DETECTION BY MULTIFRACTAL SPECTRAL ANALYSIS

Sahil Agarwal,^{1,*} Fabio Del Sordo,^{2,3,4} and John S. Wettlaufer^{1,4,5,6}

¹*Program in Applied Mathematics, Yale University, New Haven, USA*

²*Departments of Geology & Geophysics, Yale University, New Haven, USA*

³*Department of Astronomy, Yale University, New Haven, USA*

⁴*Nordita, Royal Institute of Technology and Stockholm University, SE-10691 Stockholm, Sweden*

⁵*Departments of Geology & Geophysics, Mathematics and Physics, Yale University, New Haven, USA*

⁶*Mathematical Institute, University of Oxford, Oxford, UK*

(Dated: September 21, 2016)

Owing to technological advances the number of exoplanets discovered has risen dramatically in the last few years. However, when trying to observe Earth analogs, it is often difficult to test the veracity of detection. We have developed a new approach to the analysis of exoplanetary spectral observations based on temporal multifractality, which identifies time scales that characterize planetary orbital motion around the host star. Without fitting spectral data to stellar models, we show how the planetary signal can be robustly detected from noisy data using noise amplitude as a source of information. For observation of transiting planets, combining this method with simple geometry allows us to relate the time scales obtained to primary transit and secondary exoplanet eclipse of the exoplanets. Making use of data obtained with ground-based and space-based observations we have tested our approach on HD 189733b. Moreover, we have investigated the use of this technique in measuring planetary orbital motion via doppler shift detection. Finally, we have analyzed numerical data obtained using the SOAP 2.0 tool, which simulates a stellar spectrum and the influence of the presence of a planet or a spot on that spectrum over one orbital period. We have demonstrated that, so long as the signal-to-noise-ratio ≥ 100 , our approach reconstructs the planetary orbital period, as well as the rotation period of a spot on the stellar surface.

I. INTRODUCTION

The last three decades have seen the birth of exoplanetary science. With the advent of various techniques, which include Doppler measurements, transit photometry, micro-lensing, and direct imaging, thousands of planets have been detected orbiting distant stars. A central focus is the detection of so-called *Exo-Earths*, terrestrial Earth-like planets in term of mass, orbiting around a star at a distance (*The Goldilocks Zone*) that, given sufficient atmospheric pressure, would allow for the existence of liquid water on its surface. The techniques that are most commonly used in discovering other planets are transit photometry [e.g. 1, and references therein] or Doppler measurements [e.g. 2, and references therein]. Recently, using the transit method, detection of 9 candidates for habitable planets was announced that may fall within the habitable zones of their host stars [3].

Whilst the combination of these approaches have provided an impressive range of observations, the detection of Earth analogs is a challenging problem. Indeed, the presence of instrumental and astrophysical noise are sources of uncertainty for such discoveries [e.g. 4]. The fingerprints of an exo-Earth could easily be hidden in stellar noise, or stellar signals might mimic the presence of an exoplanet. Moreover, when such noise is modeled, there is a risk of introducing spurious signals in the analysis of data [e.g. 5, 6].

Radial velocity and transit photometry methods also suffer from a high rate of false detections [3], which is in part due to the models upon which these methods are based. Such exoplanet evolution and stellar models over-fit parameters to the data, which turns out to be crucial when one aims to detect terrestrial-like planets [6, 7]. These parameters include, but are not limited to, planet and stellar radius, their masses, eccentricity, impact factor, transit duration, transit ingress/egress duration, transit depth, orbital inclination, distance of planet from the star, limb-darkening parameters (which themselves may vary on the law used), and shape of the transit curve. In some cases, instrumental signals such as the wobble of the instrument cluster aboard the Spitzer satellite or the latent charge build-up in the pixels (ramp) are also modeled [8, 9]. Additionally, noise sources such as granulation over the stellar surface employ model fitting to estimate the effect of that noise on the data [7], stellar activity concurrent with the stellar rotation are modeled by fitting sine waves through the radial velocity data, and long term stellar activity, light contamination from near by stars, among others are parametrically modeled. This fitting of models has led to the introduction of spurious signals in the observed data and hence to false detections.

In order to fit these stellar models to data, one must begin by considering a particular system, for example an unblended eclipsing binary or a transiting planet. Hence the list of models that the data can represent needs to be complete, which constitutes a weakness for these types of fitting scheme [3]. Moreover, these methods cannot always distinguish whether the signal is instrumental or from stellar activity [3]. Finally, because it affects the de-

* sahil.agarwal@yale.edu

tails of the fitting and hence the results, the observations must have a high signal-to-noise ratio.

A key aspect of the analysis of exo-planetary systems is the identification of periodic timescales. This is also central in the study of multi-planetary systems [e.g. 10], and often requires the combined use of the two most successful approaches in exoplanet searches; the transit and radial velocity methods. The most widely used method for identifying periodic timescales is the Lomb-Scargle periodogram, which is based on the assumption that the data can be interpreted as a sum of periodic signals. However, this technique is used on data that have been “filtered” using models for stellar noise and terrestrial atmospheric contamination. Hence, it is again possible to accumulate artifacts in the data through the use of such models.

Here, in order to extract the timescales that characterize a planetary system, we introduce a new approach to the analysis of spectral data that makes no assumptions about the data itself. In the spirit of the Langevin theory of Brownian Motion, we quantify a signal coming from a star as the combination of a deterministic dynamics and stochastic noise, but make no *a priori* assumptions about the nature of these dynamical processes. Thus, we examine an unfiltered time series X_i of a stellar spectral signal. If the star hosts a planetary system, the time scales associated with stellar rotation and activity, as well as with planetary motions must be present. We need not (a) make assumptions regarding the combination of periodic signals, and (b) use stellar models. The goal is to agnostically identify the dominant timescales of the observed system, and then use elementary geometry to reveal the underlying dynamics of the system.

We begin with a description of the method in §II. The proposed approach can be used both for transiting planets and for radial velocity measurements. We show two examples in §III; one for a transiting planet, and the other for a simulated observation of a planet detectable only via radial velocity measurements. Finally, we discuss the results and their robustness in §IV, and conclude in §V.

II. METHOD

We analyze a series of spectra taken at an approximately constant time intervals. If each of these spectra spans a wavelength range of L wavelengths, we construct L time series, each of which consists of the flux measured at a given wavelength. Hence, if we have m equally-spaced-in-time spectra, we have L time series of length m . The time delay between spectral observations corresponds to the best obtainable resolution.

At each wavelength the variability of flux in time can arise from, among other things, the doppler shift, photometric effects, atmospheric/telluric effects, and instrumental noise, with characteristic time scales we aim to extract with our method. Importantly, with no *a priori* knowledge of the dynamics of the system and without

the use of model-fitting, we can extract the time scales associated with either a transiting planet or the doppler shift underlying planetary motion. In the former case we use elementary geometry to reconstruct time scales connected with the transit. In the latter case it is known that a doppler shift for stellar spectra can be caused both by an orbiting planet as well as by intrinsic stellar features, such as spots. We shall show that we can identify both of them, but to distinguish between them is the subject of future work.

A. Multi-fractal Temporally Weighted Detrended Fluctuation Analysis

We analyze spectral time series using Multi-fractal Temporally Weighted Detrended Fluctuation Analysis (MF-TW-DFA) [e.g., 11, and references herein], which does not *a priori* assume anything about the temporal structure of the data. The approach has four stages, which we describe briefly below.

- We construct a non-stationary *profile* $Y(i)$ of the original time series X_i as,

$$Y(i) \equiv \sum_{k=1}^i (X_k - \bar{X}_k), \quad \text{where} \quad i = 1, \dots, N. \quad (1)$$

The profile is the cumulative sum of the time series and \bar{X}_k is the average of the time series $X_1 \dots X_k$.

- This non-stationary profile is divided into $N_s = \text{int}(N/s)$ non-overlapping segments of equal length s . The time series has a length that is rarely an exact multiple of s , which is handled by repeating the procedure from the end of the profile and returning to the beginning, thereby creating $2N_s$ segments.
- A point by point approximation $\hat{y}_\nu(i)$ of the profile is made using a moving window, smaller than s and weighted by separation between the points j to the point i in the time series such that $|i - j| \leq s$. [12] A larger (or smaller) weight w_{ij} is given to $\hat{y}_\nu(i)$ according to whether $|i - j|$ is small (large) [11]. This approximated profile is then used to compute the variance spanning up ($\nu = 1, \dots, N_s$) and down ($\nu = N_s + 1, \dots, 2N_s$) the profile as

$$\begin{aligned} \text{Var}(\nu, s) &\equiv \frac{1}{s} \sum_{i=1}^s \{Y([\nu - 1]s + i) - \hat{y}([\nu - 1]s + i)\}^2 \\ &\quad \text{for } \nu = 1, \dots, N_s \quad \text{and} \\ \text{Var}(\nu, s) &\equiv \frac{1}{s} \sum_{i=1}^s \{Y(N - [\nu - N_s]s + i) - \hat{y}(N - [\nu - N_s]s + i)\}^2 \\ &\quad \text{for } \nu = N_s + 1, \dots, 2N_s. \end{aligned} \quad (2)$$

- Finally, a *generalized fluctuation function* is obtained and written as

$$F_q(s) \equiv \left[\frac{1}{2N_s} \sum_{\nu=1}^{2N_s} \{\text{Var}(\nu, s)\}^{q/2} \right]^{1/q}. \quad (3)$$

The behavior of $F_q(s)$ depends on the choice of time segment s for a given order q of the moment taken. The principal focus is to study the scaling of $F_q(s)$ as characterized by a generalized Hurst exponent $h(q)$ viz.,

$$F_q(s) \propto s^{h(q)}. \quad (4)$$

When $h(q)$ is independent of q the time series is said to be monofractal, in which case $h(q)$ is equivalent to the classical Hurst exponent H . For $q = 2$, regular MF-DFA and DFA are equivalent [13], and $h(2)$ can also be related to the decay of the power spectrum S . If $S(f) \propto f^{-\beta}$, with frequency f then $h(2) = (1 + \beta)/2$ [e.g., 14]. For white noise $\beta = 0$ and hence $h(2) = 1/2$, whereas for Brownian or red noise $\beta = 2$ and hence $h(2) = 3/2$. The dominant timescales in the data set are the points where the fluctuation function $\log_{10} F_2(s)$ changes slope with respect to $\log_{10} s$.

III. DEMONSTRATING THE METHOD WITH THREE TYPES OF DATA

To test the method described in §II we apply it to three different kinds of data. The first two are spectral observations of HD 189733b, a well known transiting exoplanet discovered in 2005 [15]. The third is a set of simulated stellar spectra affected by the doppler shift induced by an orbiting planet. These simulated data are obtained with the Spot Oscillation And Planet (SOAP) 2.0 tool [16].

A. HD 189733b

First detected in 2005 [15], HD 189733b is a hot Jupiter orbiting the star HD 189733A in the constellation Vulpecula, approximately 63 light years away from Earth. We use spectral data for HD 189733b to test our method of extracting time scales related to the effects of a transiting planet. We employ both the ground-based observations (high-resolution in wavelength) and space-based observations (low-resolution in wavelength) to test if and how the analysis is affected by resolution, instrumental noise, and the terrestrial atmosphere. The ground-based observations are obtained from the High Accuracy Radial velocity Planet Searcher *HARPS* spectrograph [17], and the space-based observations are from the Nasa Spitzer space mission.

1. HARPS Data

These are reduced 1-D spectral data from the High Accuracy Radial velocity Planet Searcher at the ESO La Silla 3.6m telescope for planet HD 189733b. Programs 072.C-0488(E), 079.C-0127(A) and 079.C-0828(A) are used from the European Southern Observatory (ESO)

TABLE I. **HARPS Data Analyzed for HD 189733b**

	Observation Date	Program ID
Night 1	2006-09-07	072.C-0488(E)
Night 2	2007-07-19	079.C-0828(A)
Night 3	2007-08-28	079.C-0127(A)

TABLE II. **Spitzer Data Analyzed for HD 189733b**

AOR Key	Observation Date	Wavelength Range (μm)
18245632	2006-10-21	7.4-14.0
20645376	2006-11-21	7.4-14.0
23437824	2008-05-24	7.4-14.0
23438080	2008-05-26	7.4-14.0
23438336	2008-06-02	7.4-14.0
23438592	2008-05-31	7.4-14.0
23438848	2007-10-31	7.4-14.0
23439104	2007-11-02	7.4-14.0
23439360	2007-06-26	7.4-14.0
23439616	2007-06-22	7.4-14.0
23440384	2008-06-09	5.0-7.5
23440640	2008-06-04	5.0-7.5
23440896	2007-12-07	5.0-7.5
23441152	2007-11-06	5.0-7.5
23441408	2007-11-11	5.0-7.5
23441664	2007-11-09	5.0-7.5
23441920	2007-11-24	5.0-7.5
23442176	2007-11-15	5.0-7.5
23439872	2007-11-04	13.9-21.3
23440128	2007-06-17	13.9-21.3
23442432	2007-12-10	19.9-39.9
23442688	2007-06-20	19.9-39.9

archive (see Table I). These data cover four nights, but that from the fourth night has been removed from this analysis as is known to have been affected by severe weather [18, 19]. Each night is treated as an independent data set. On night 1, these spectra are taken approximately every 10.5 minutes, and on nights 2 and 3 every 5.5 minutes.

2. Spitzer Data

The Spitzer space mission observed secondary eclipses of HD 189733b. We use Basic Calibrated Data (.bcd) files from the Spitzer pipeline version 18.18.0, obtained with the optimal extraction tool in the SPICE software that employs the Optimal Extraction Algorithm of Horne [20] to obtain the reduced 1-D spectra from the observed images (see Table II).

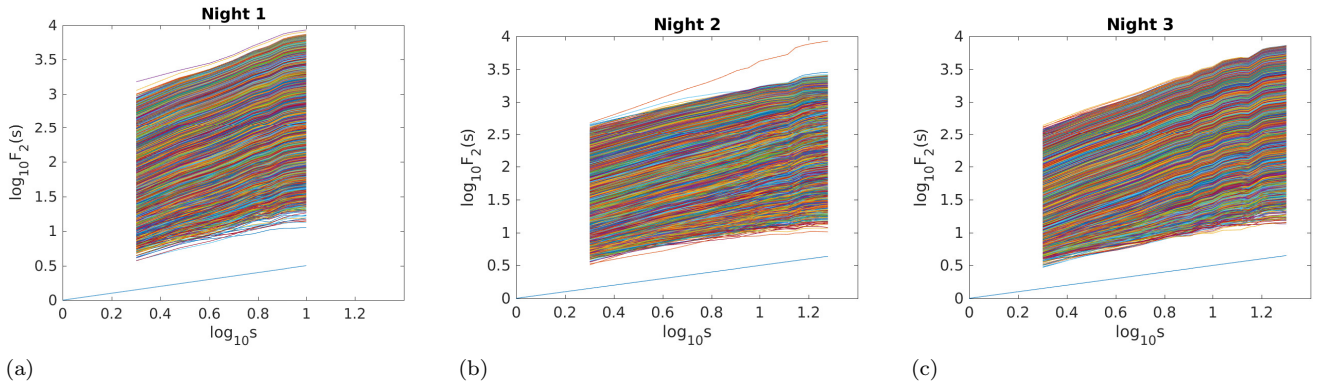


FIG. 1. The second moment of the fluctuation functions for (a) Night 1 (time resolution 10.602988 ± 0.072236 minutes), (b) Night 2 (time resolution 5.517202 ± 0.005683 minutes) and (c) Night 3 (time resolution 5.523740 ± 0.004415 minutes), with original wavelength resolution (plotted at alternate wavelengths). Different colors represent different wavelengths. The straight blue line has a slope of 0.5, the slope of fluctuation function denoting white noise. Shorter wavelengths show higher amounts of noise which may be associated with atmospheric turbulence and/or telluric effects on Earth or the exoplanet.

B. Simulations

The use of simulated data provides a good test and application of our method to radial velocity measurements, because we can control the input and thereby facilitate a clear interpretation of the results. We use data produced with the SOAP 2.0 tool [16]. These spectra are generated by filling a grid of cells with the solar spectrum from National Solar Observatory, simulating a rotating star. Depending on the grid cell position, the spectrum is Doppler shifted to account for stellar rotation, and the intensity of each cell is weighted by a limb-darkening law. To simulate the presence of a spot, a typical spectrum of a solar spot is enclosed in a cell. Additionally, the intensity is reduced in the presence of a dark spot or increased in the presence of a bright plage. Such features follow the rotation of the star. The final integrated spectrum is the sum of the spectra from all of the cells.

First, we use stellar spectra in the absence of stellar spots and thus influenced solely by an orbiting planet. This allows us to focus on, for example, the “bare” effect of the Doppler shift on the spectra. Because we would like to examine orbital periods, we stack the 25 spectra corresponding to the orbital period of the planet to provide 8 orbital periods worth of data. Moreover, since we are interested in the effect of noise on our data analysis, Gaussian white noise with a specific signal to noise ratio (SNR) is then added at each time to obtain data sets with different SNRs. This is important for the comparison of results obtained with this simulated dataset to real observations from spectrographs characterized by different SNRs.

Second, we use stellar spectra in the absence of a planet and thus influenced solely by a stellar spot. This allows us to focus on the “bare” effect of a spot, which can in principle mimic the Doppler shift in radial velocity measurements induced by a spot. Here again as we have 25 spectra corresponding to one full rotation period of

the star, we stack these to give 8 orbital periods worth of data, and then add Gaussian white noise as in the first (planet only) case discussed above.

IV. DISCUSSION

A. HARPS analysis of HD 189733b

The 1-D spectra from the HARPS instrument provides the time series for flux at each of the wavelengths for each night separately. These are analyzed using Multi-fractal Temporally Weighted Detrended Fluctuation Analysis as described in §II above. The second moment of the associated fluctuation functions are shown in Figures 1a, 1b and 1c for Nights 1, 2 and 3 respectively.

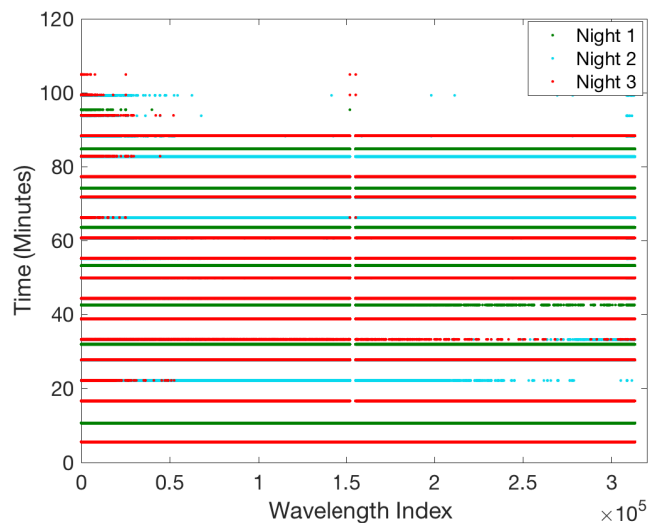


FIG. 2. The crossover times plotted for all wavelengths, for all three nights for HD 189733b. Only the 85 minute timescale is present for all three nights.

The key aspects of the fluctuation functions are as follows. First, nearly all of these curves are parallel to each other, demonstrating that the spectra evolve in time with a similar noisy behavior at all wavelengths. Second, those fluctuation functions that show deviations, do so principally at smaller wavelengths and can thus be ascribed to atmospheric interference from the Earth or the exoplanet, such as variability of air masses, telluric contamination and/or turbulent effects. Third, for all nights and all the wavelengths we see a timescale of 85 minutes. Fourth, for times larger than 85 minutes, the dynamics for all wavelengths exhibit a white noise structure.

In Figure 2 we plot all of the times at which the slope of the fluctuation functions changes—the crossover times—for the three nights for all of the wavelengths. The robustness of the 85 minute time scale is reflected in by its presence for all nights, whereas other times scales are present for only one or two nights. Possible origins of the other time scales include turbulent and convective processes in Earth’s and the exoplanet’s atmosphere, stellar activity or instrumental noise. It is important that we have identified these scales here and this provides a foundation for systematic examination of them in future studies as a possible means of systematically filtering them.

To understand this 85 minute time scale, we study the following parameters of the planet HD 189733b, as found from previous studies: (1) the ratio of the planetary (R_p) to the stellar (R_s) radii, $(R_p/R_s)^2 = 0.02391 \pm 0.00007$ [21], and (2) the duration of the transit of the planet in front of the star $\tau_{14} = 0.07527 \pm 0.00037$ days ≈ 108.4 minutes, or the time between first and last “contact” of the planet and the star [18] (see Figure 3).

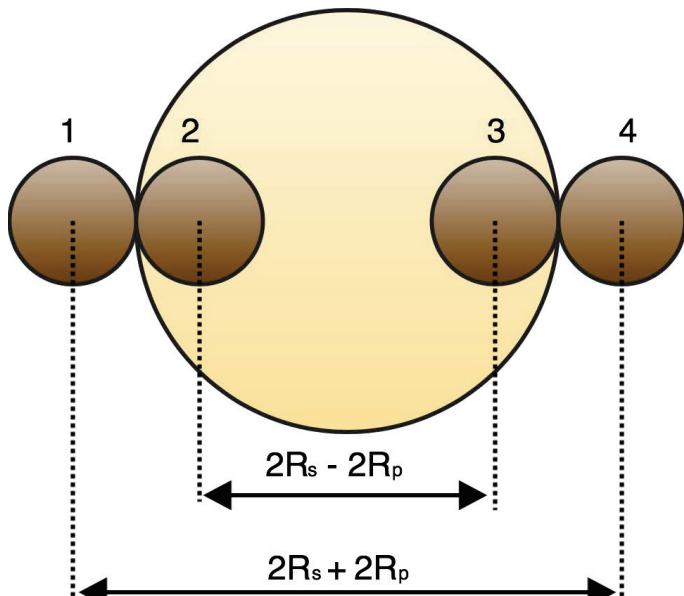


FIG. 3. A schematic showing the planet of radius R_p transiting its parent star of radius R_s .

Due to the time-resolution of the data for each night, as well as the total length of these time series, we are not

able to observe the shortest and the longest time scales in the system. The 85 minute time scale corresponds to the interval between the second and third contact of the planet, τ_{23} , i.e., the period during which the planet is completely between the observer and the star (Figure 3). If we assume that the planet is transiting in the equatorial plane of the star, and hence the impact parameter is $b = 0$, we calculate this time by tracking the center of the planetary disk across the stellar disk as

$$\frac{\tau_{14}}{2(R_s + R_p)} = \frac{\tau_{23}}{2(R_s - R_p)}, \quad (5)$$

from which we obtain $\tau_{23} \approx 0.0551$ days = 79.659 minutes, using $\frac{R_p}{R_s}$ from Torres *et al.* [21]. Within the resolution of the time series, this value of τ_{23} is consistent with the 85 minute time scale we have found. In the next section, we provide a more detailed explanation for this time scale.

To test how wavelength-resolution affects our results, we degrade the resolution of the spectra in wavelength and repeat our analysis. We decrease the resolution by a factor 100 by using the average of each 100 points, and plot the fluctuation functions for the degraded spectra from Night 1 in Figure 4. Clearly, the structure at longer time scales is unchanged, whereas the noisy behavior at short wavelengths is less evident than in the analysis of high resolution spectra, showing that the averaging acts as a crude high-pass filter. Importantly, the 85 minute scale persists as does the general behavior of the fluctuation functions for this reduction in resolution.

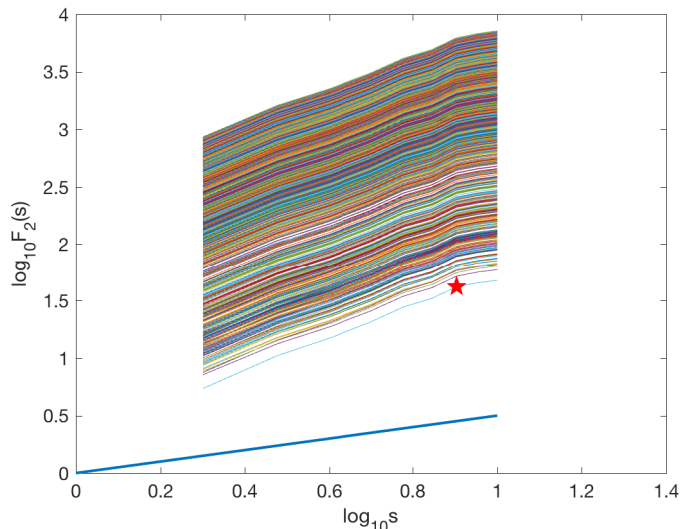


FIG. 4. The second moment of the fluctuation functions for Night 1 (approximate time resolution 10.5 minutes), with a wavelength resolution degraded of a factor 100. As in figure 2, the straight blue line has a slope of 0.5, the slope of the fluctuation function denoting white noise.

B. Spitzer analysis of HD 189733b

Because stars can only be observed at night, all Earth based instruments provide a strong constraint on the detection of time scales. Moreover, telluric contamination is a well known problem that is also evident in our method at the shorter timescales in the HARPS data. To bypass these limitations, we examine data from the Spitzer mission, which observed HD189733b for 22 nights, looking at the secondary eclipses of the planet, i.e., when the planet is behind the star. Although the data from Spitzer is low resolution in the wavelength space, as we have just shown in the case of HARPS data this only influences the short time scales, and does not affect the identification of longer and robust time scales. Regardless of the wavelength resolution, the time resolution as well as the total duration make the Spitzer datasets compelling.

As was done in Figure 2, we plot these crossovers as a function of wavelength in Figure 5. It is clear that relative to the HARPS data the Spitzer data is substantially more noisy. We note that typically the raw data are filtered to remove the noisy characteristics, but given experience with other systems [11, 22] we take the perspective that noise can be an essential source of information. The four prominent timescales in this data are (1) $\tau_{12} =$

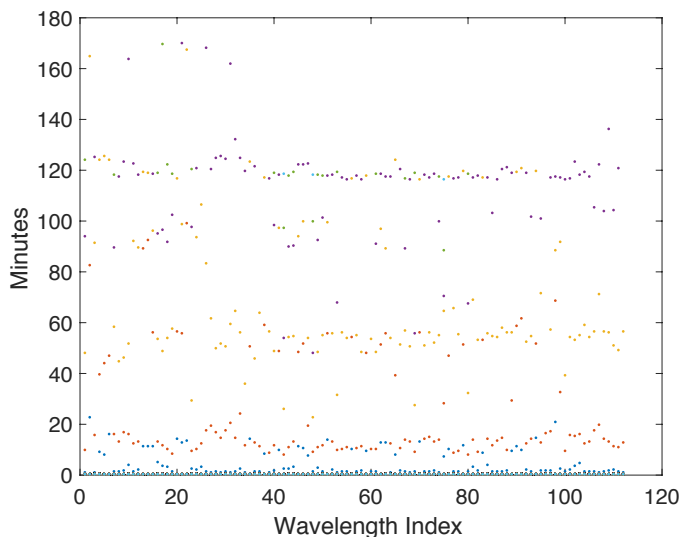


FIG. 5. The Spitzer based crossover times plotted for all wavelengths, for Night 2 (AOR-20645376, see Table II) for HD 189733b. All four significant timescales are robustly extracted using our method.

15.4804 ± 3.7660 minutes, (2) 55.0966 ± 5.8851 minutes, (3) $\tau_{23} = 87.0947 \pm 2.7888$ minutes, and (4) $\tau_{14} = 118.9671 \pm 5.5764$ minutes, where the uncertainty is one standard deviation about the mean. First, the 55.0966 ± 5.8851 minute timescale is the pointing wobble in the Infrared Array Camera of the Spitzer telescope, which is due to the battery heater in the telescope [8, 23]. Secondly, the 85 minute timescale obtained from the HARPS data is within one standard deviation of τ_{23} and hence is

robust. The other two timescales are related to the transit of the planet behind the star, namely the secondary eclipse. The necessary and sufficient condition for these timescales to represent the transit (Figure 3) is

$$\tau_{14} = \tau_{23} + 2 \times \tau_{12} \quad (6)$$

It is evident that, within one standard deviation, equation 6 is satisfied for the above timescales.

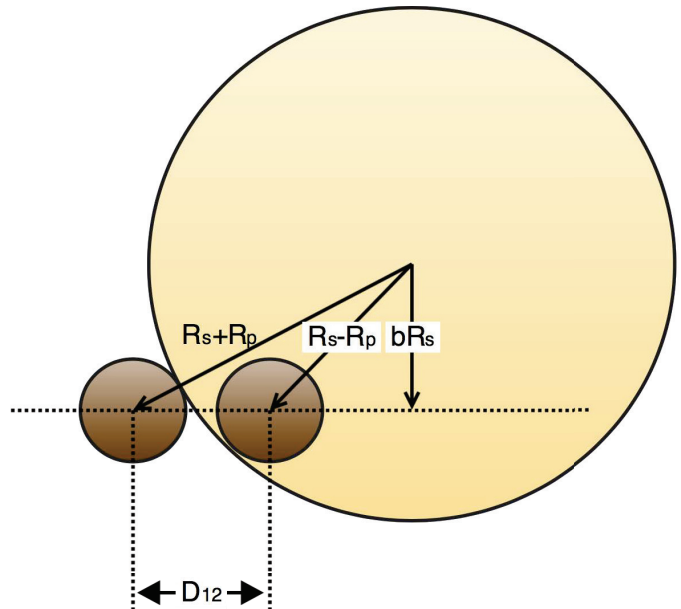


FIG. 6. Schematic showing a planet transiting its parent star at any general latitude measured by the impact parameter b .

Using the values for τ_{14} and τ_{23} , we can then estimate the ratio of the radius of the planet to that of the star by rearranging Equation 5 to obtain

$$\frac{R_p}{R_s} = \frac{\tau_{14} - \tau_{23}}{\tau_{14} + \tau_{23}} = 0.1543 \pm 0.0283. \quad (7)$$

Therefore, we have shown here that without the use of any fitting of model parameters such as the epoch of mid-transit, orbital period, fractional flux deficit, total duration of transit, the impact parameter of planet's path across the stellar disc, the transit depth, the shape parameter for transit, the transit ingress/egress times and many more [3, 24], we can now calculate most of these from the time scales alone [25]. We note that, to within the precision of ± 0.0283 , the result in equation 7 is the same as that found by [15] and the average of that found by [26].

In this analysis we did not take into account the possibility that the planet may be transiting with a non-zero impact parameter ($b \neq 0$, where $b = 0$ implies the planet is transiting the equator of the star from the observer's perspective). This would add another unknown in equation 7. Nevertheless simple geometry as shown in Figure 6 provides another equation, also involving τ_{12} , which gives a system of two second order equations that can

be solved to determine both the ratio of radii and the impact parameter as follows.

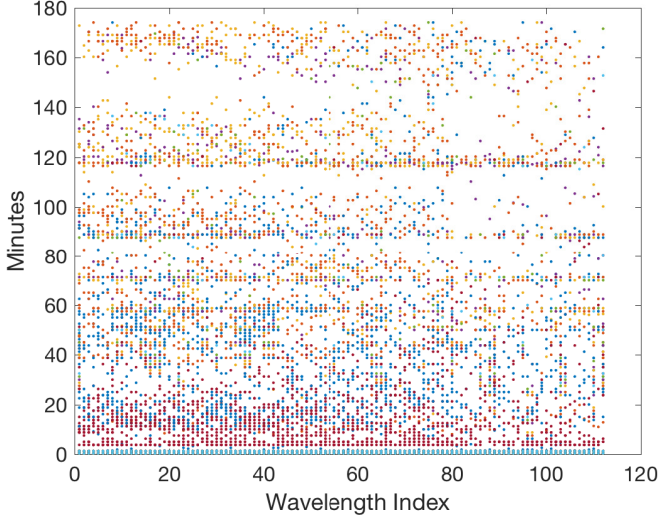


FIG. 7. The crossover times plotted for all wavelengths, for those nights which have high SNR [9] for HD 189733b. All four significant timescales are robustly extracted using our method.

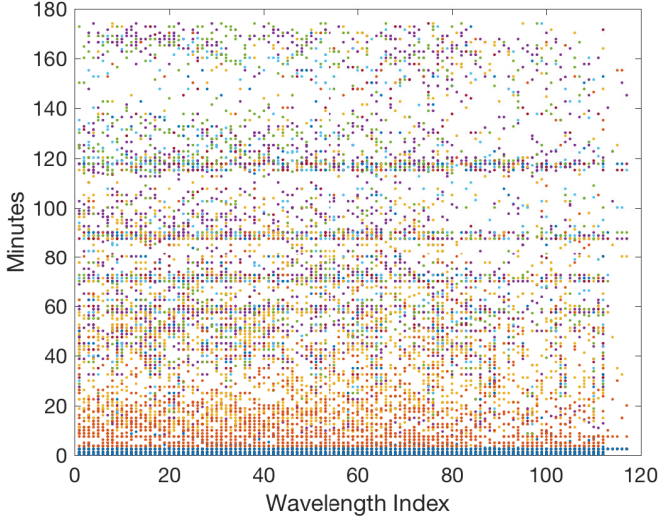


FIG. 8. The crossover times plotted for all wavelengths, for all nights for HD 189733b. All four significant timescales are robustly extracted using our method.

First we can write $D_{12} = D_{34}$, the distance between the first-second or third-fourth contact from the Pythagorean theorem as

$$D_{12} = D_{34} = \sqrt{(R_s + R_p)^2 - b^2 R_s^2} - \sqrt{(R_s - R_p)^2 - b^2 R_s^2}. \quad (8)$$

Second, D_{23} , the distance between the second-third contact is

$$D_{23} = 2\sqrt{(R_s - R_p)^2 - b^2 R_s^2}, \quad (9)$$

and finally we have

$$D_{14} = 2\sqrt{(R_s + R_p)^2 - b^2 R_s^2}. \quad (10)$$

Therefore,

$$\frac{\tau_{14}}{D_{14}} = \frac{\tau_{23}}{D_{23}}, \quad (11)$$

which, upon substitution of Eqs 9 and 10, gives

$$\tau_{23} = \tau_{14} \times \sqrt{\frac{1 + \rho^2 - 2\rho - b^2}{1 + \rho^2 + 2\rho - b^2}}, \quad (12)$$

where $\rho = \frac{R_p}{R_s}$. Similarly, Eqs 8 and 10 give

$$\tau_{12} = \tau_{14} \times \frac{1}{2} \left(1 - \sqrt{\frac{1 + \rho^2 - 2\rho - b^2}{1 + \rho^2 + 2\rho - b^2}} \right), \quad (13)$$

and thus Eqs. 12 and 13 can be used to calculate ρ and b simultaneously. As a check, Eqs. 12 and 13 combine to give Eq. 5, and Eq. 12 becomes Eq. 7 in the case $b = 0$.

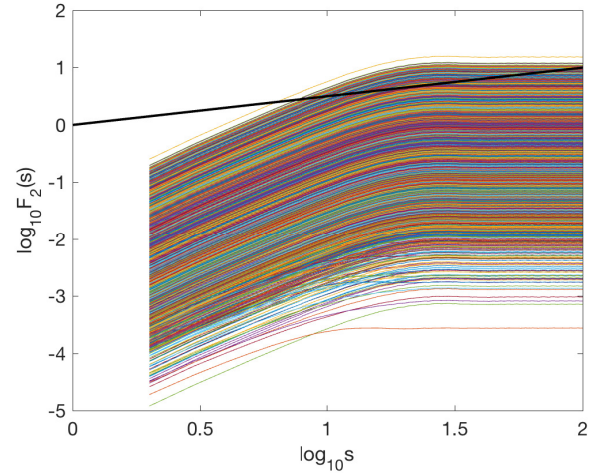


FIG. 9. The second moment of the fluctuation functions are shown for all the wavelengths for the simulated planet without noise in the spectra. The straight black line has a slope of 0.5, which denotes white noise dynamics.

In Figures 7 and 8 we show the crossovers from all the available datasets, as shown in Table II. Additionally, Figure 8 includes those data sets that are not used in other studies due to low signal to noise ratio, or other issues [9]. Moreover, we see the clear emergence of the significant time scales discussed above from *all* of these data sets. Importantly we note the amount of noise in these data sets, which as we will see below can be a source of information for the robust estimation of these time scales.

C. SOAP Simulated Data

Thus far we have focused our analysis on data observed during a primary or secondary eclipse of an exoplanet.

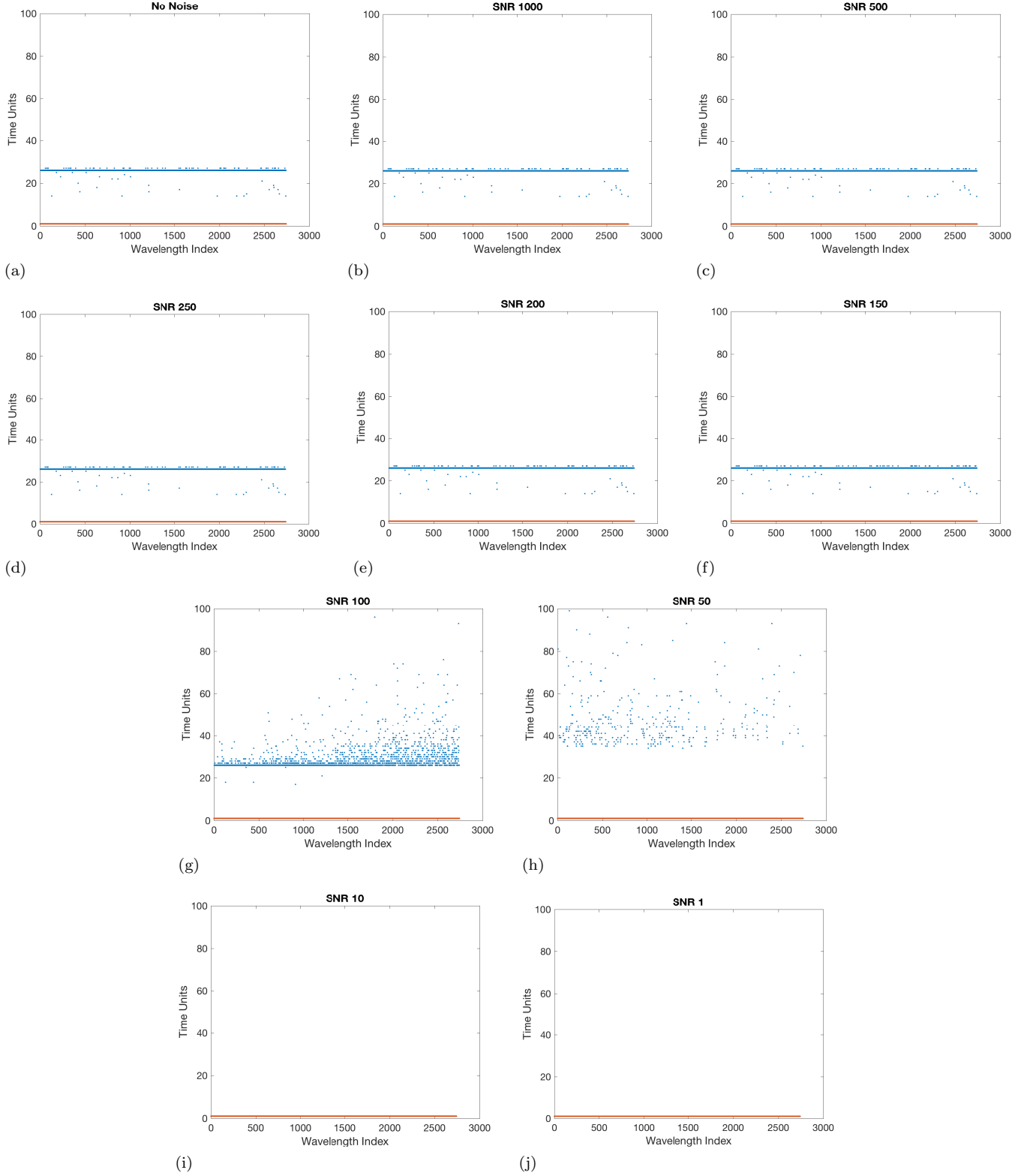


FIG. 10. The crossover times plotted for all wavelengths, with $C_{th} = 0.08$, for the simulated SOAP 2.0 spectra in presence of an orbiting planet. We show results for different SNR. (a) No Noise, (b) SNR = 1000, (c) SNR = 500, (d) SNR = 250, (e) SNR = 200, (f) SNR = 150, (g) SNR = 100, (g) SNR = 50, (i) SNR = 10, (j) SNR = 1.

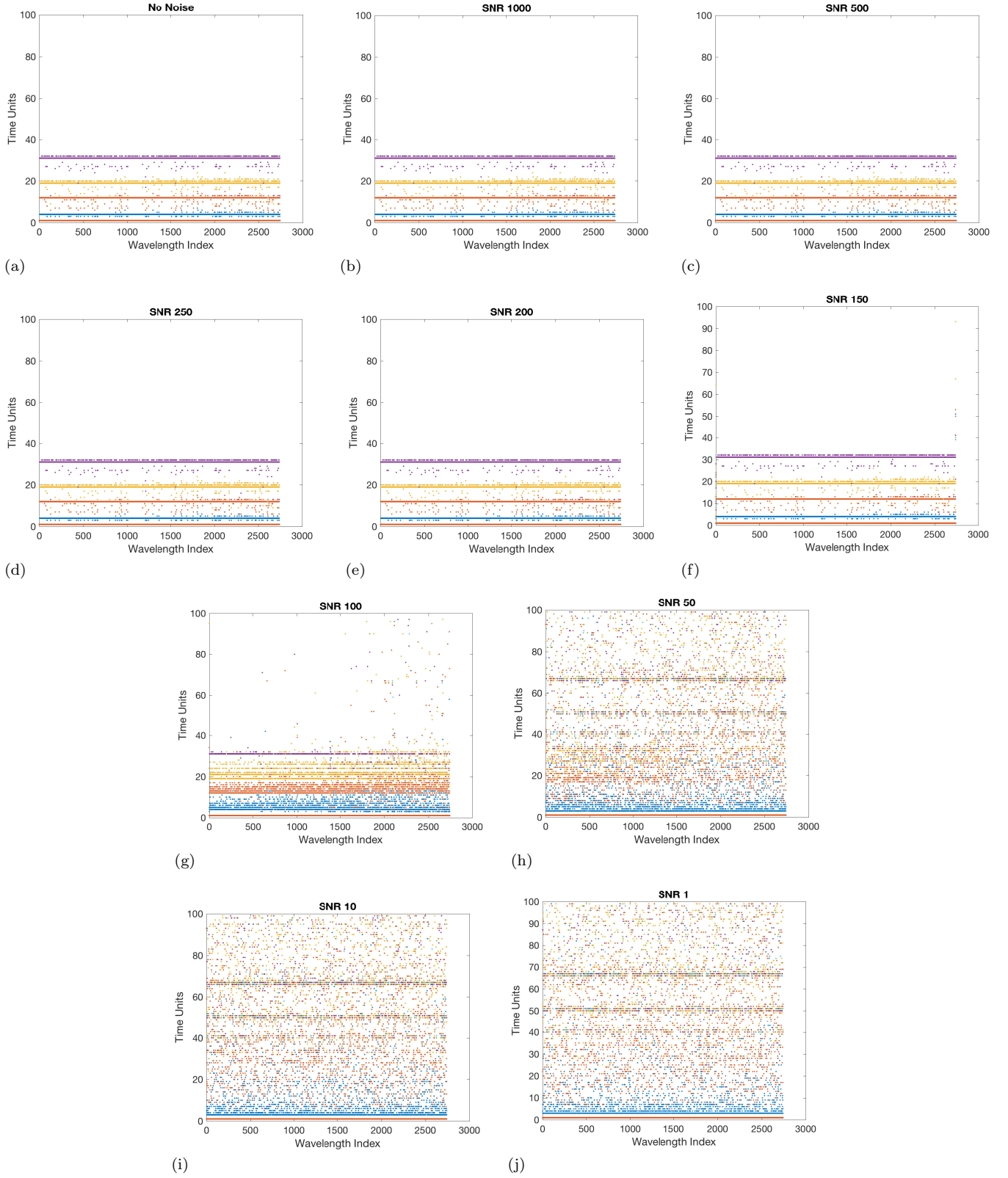


FIG. 11. The crossover times plotted for all wavelengths, with $C_{th} = 0.01$, for the simulated SOAP 2.0 spectra in presence of an orbiting planet. We show results for different SNR. (a) No Noise, (b) SNR = 1000, (c) SNR = 500, (d) SNR = 250, (e) SNR = 200, (f) SNR = 150, (g) SNR = 100, (g) SNR = 50, (i) SNR = 10, (j) SNR = 1.

However, these measurements are not always available since only a small percentage of exoplanet transits occur between the observer and the host star. Instead, all the exoplanets produce an effect on the radial velocity of the host star.

Current technology only allows detection of the red and blue shifts of the spectrum when the motion of the star is sufficiently large, for example when a Jupiter-sized planet orbits very closely its parent star. The detection of a smaller planet is thus often controversial. Moreover, these data are contaminated with a wide variety of noise, such as instrumental noise, atmospheric noise, stellar noise, absorption lines. Presently, the observations are fit to a radial velocity curve, which may or may not be robust and have a non-negligible false detection rate [see sect 4.3.2. of 4].

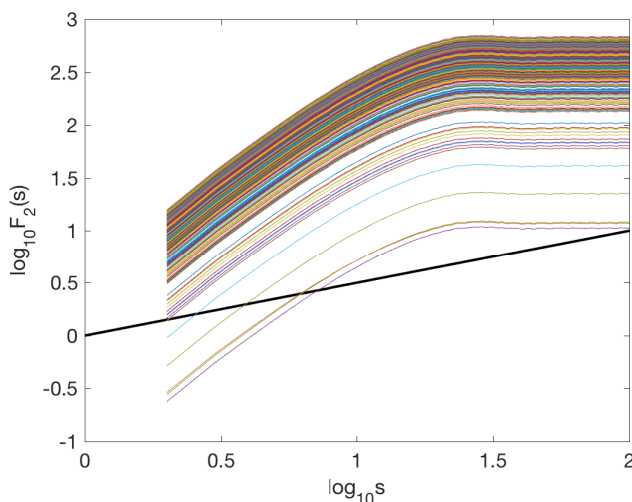


FIG. 12. The second moment of the fluctuation functions for all wavelengths for the simulated SOAP 2.0 spectra with 5% of the stellar disk covered with a spot, without noise in the spectra. The straight black line has a slope of 0.5, which denotes white noise dynamics.

To address this problem, we test the detectability of time scales related to the radial velocity effect with our method. We start from the analysis of a well defined case using data from the SOAP 2.0 simulations. This allows us to study the effect of noise as well as the robust estimation of crossover times, which we know to be present in the data set.

The original (“noise-free”) data describe the shifts in the spectrum of the star with a radial velocity of amplitude 40 ms^{-1} , due to a planet orbiting around it. We add Gaussian white noise to the spectra at each time to get a specific SNR, and thereby obtain 10 such data sets for SNR 1, 10, 50, 100, 150, 200, 250, 500 and 1000. Figure 9 shows the second moment of the fluctuation functions for all the wavelengths in the reduced spectra without noise, and subfigures in Figure 10 show the crossover time scales extracted for the different noise cases. To obtain these crossovers, we calculate the change in slope

of the fluctuation functions for each wavelength, based on a threshold value. Namely, a crossover is calculated if the change in slope of the curve is more than the set threshold, C_{th} and the plots in Figure 10 use a threshold value $C_{th} = 0.08$. Whence, we are able to extract the exact timescale for the orbital period of the exoplanet. Even for a $SNR = 150$, the methodology is efficient and robust against noise, and we extract the correct timescale. As the signal quality is degraded further, we see the noise starting to affect the calculated crossovers with scatter above the actual value. Finally, as we further degrade the signal quality the crossover times disappear altogether. As the noise begins to dominate, the signal becomes white noise. We know this quantitatively for in our analysis fluctuation functions with a constant slope of 0.5 are by definition white noise processes. To examine how to capture timescales when the signal quality is poor, we decrease the threshold to $C_{th} = 0.01$ (Figure 11). Even this threshold is able to approximately capture (31 time units) the orbital period, but it also detects other multiple time scales, which may be spurious. The important point here is the ability to study how noise affects these timescales; as the SNR is decreased time scales remain robust, but become harmonics of the robust scales for further increase in the SNR. These may be due to a form of stochastic resonance between the threshold, radial velocity measurements and noise. The value of threshold is thus crucial.

In actual data, one cannot calculate this threshold for all the wavelengths separately and for each night, and hence one value is chosen in accordance with the observed noise characteristics. Presently, use of periodograms and fitting to sine curves is done to model the radial velocity curves from the observations. The above analysis shows how noise can lead to spurious estimation of orbital periods thus potentially result in spurious detection of exoplanets.

In the same manner as above, we have analyzed the case in which no planet is orbiting around the star, but a spot covers 5% of its surface. In a similar fashion we have analyzed 8 rotation periods of the star and reconstructed its rotational period (Figure 12, 13 and 14). The use of this method to distinguish stellar features from planets will be subject of future work.

V. CONCLUSION

We have presented and tested a new multi-fractal approach to the analysis of exo-planetary spectral observations. The goal is to use a fit-free procedure to identify robust time scales associated with the exo-planetary orbital motion around a host star. With these timescales in hand, one can compute key system parameters such as the ratio of the size of the planet to that of the star and latitude of transit, *without* use of stellar evolution models, data fitting, noise filtering, and the additional wide variety of other assumptions about the system that are

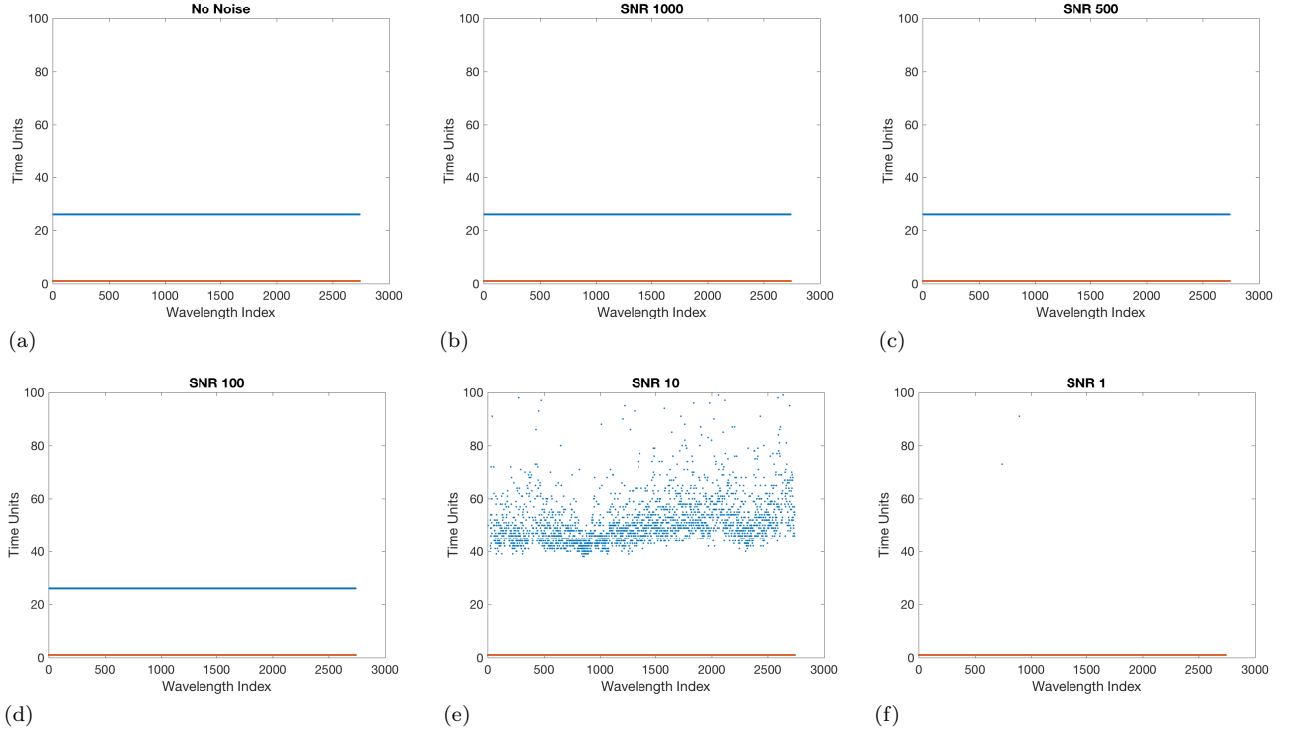


FIG. 13. The crossover times plotted for all wavelengths, with $C_{th} = 0.08$, for the simulated SOAP 2.0 spectra with 5% of the stellar disk covered with a spot, with different SNR. (a) No Noise, (b) SNR = 1000, (c) SNR = 500, (d) SNR = 100, (e) SNR = 10, (f) SNR = 1.

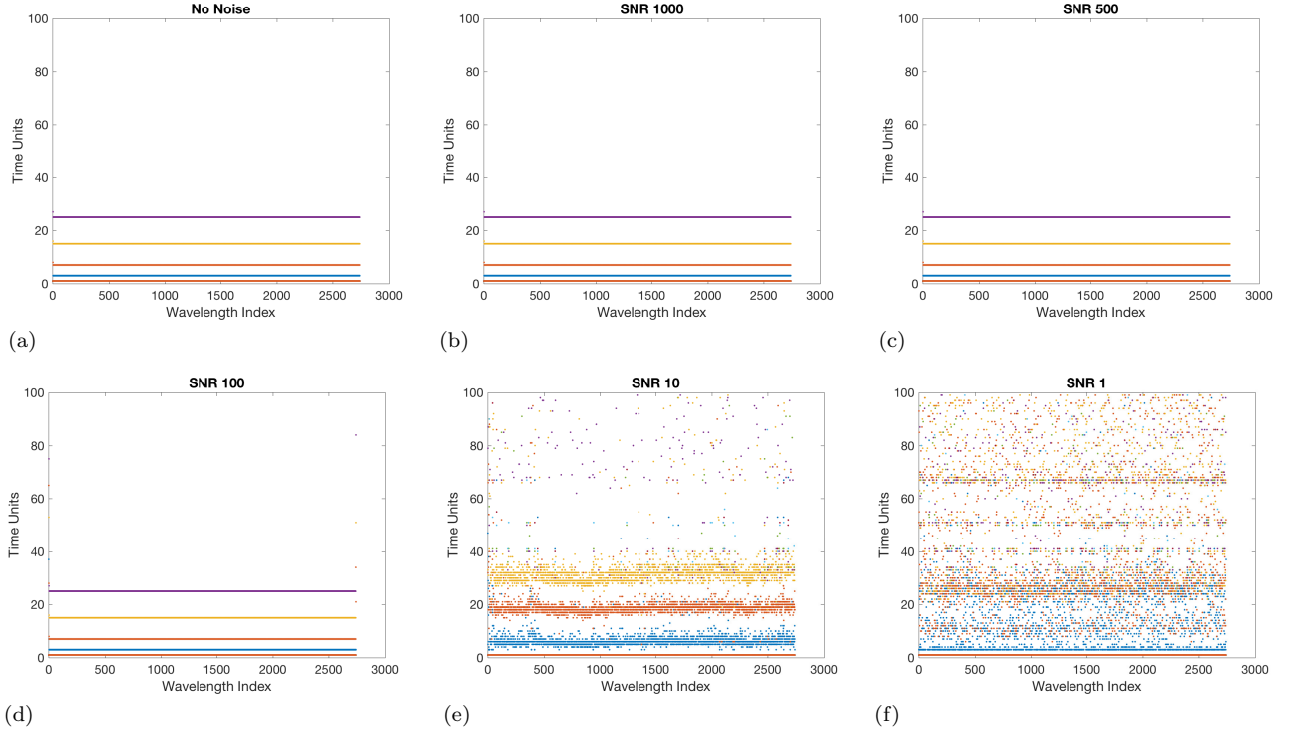


FIG. 14. The crossover times plotted for all wavelengths, with $C_{th} = 0.01$, for the simulated SOAP 2.0 spectra with 5% of the stellar disk covered with the spot, with different SNR. (a) No Noise, (b) SNR = 1000, (c) SNR = 500, (d) SNR = 100, (e) SNR = 10, (f) SNR = 1.

typically made. The concept of the approach is to take an agnostic (or model free) view of the observed spectral structure. The method makes no *a priori* assumptions about the temporal structure in any observed spectra and makes use of only one number, the generalized Hurst exponent, the value of which underlies the identification of the key time scales. We have reconstructed the primary and secondary transit times of the exoplanet HD 189733b, using data from both ground-based (HARPS spectrograph) and space-based (Spitzer) observations.

Using the SOAP 2.0 tool, which simulates a stellar spectrum and how it can be affected by the presence of a planet across one orbital period, our method is further tested in the context of measuring planetary orbital motion via doppler shift detection. Because the SOAP tool can also simulate the presence of a spot on the stellar surface across one rotational period of the star, we have demonstrated that our approach reconstructs the planetary orbital period, as well as the rotation period of a spot covering 5% of the stellar surface. Moreover, we have tested the analysis with a wide range of signal to noise ratios. We find that provided the $\text{SNR} \geq 100$ we can reconstruct (a) the period of a planet producing a 40 m s^{-1} doppler shift of the stellar spectrum, and (b) the rotational period of the star, based on the presence

of a spot on its surface. Importantly, to avoid introduction of errors arising from intrinsic irregularities in the analyzed time series, this method has the highest fidelity when observations have a regular high frequency.

In conclusion this method based on Multi-fractal Temporally Weighted Detrended Fluctuation Analysis of time series is a robust way to measure planetary orbital motion. It provides a fertile framework to examine other data sets and to explore trying to systematically distinguish stellar noise from planetary motion.

ACKNOWLEDGMENTS

The authors thank Xavier Dumusque for providing the SOAP 2.0 data used for this work and Debra Fischer for discussions on this and other detection schemes. SA and JSW acknowledge NASA Grant NNH13ZDA001N-CRYO for support. FDS is supported by the Swedish Research Council International Postdoc fellowship. JSW acknowledges Swedish Research Council grant no. 638-2013-9243 and a Royal Society Wolfson Research Merit Award for support.

-
- [1] J. J. Lissauer, R. I. Dawson, and S. Tremaine, *Nature* **513**, 336 (2014).
 - [2] M. Mayor, C. Lovis, and N. C. Santos, *Nature* **513**, 328 (2014).
 - [3] T. D. Morton, S. T. Bryson, J. L. Coughlin, J. F. Rowe, G. Ravichandran, E. A. Petigura, M. R. Haas, and N. M. Batalha, *ApJ* **822**, 86 (2016).
 - [4] D. A. Fischer, G. Anglada-Escude, P. Arriagada, R. V. Baluev, J. L. Bean, F. Bouchy, L. A. Buchhave, T. Carroll, A. Chakraborty, J. R. Crepp, R. I. Dawson, S. A. Diddams, X. Dumusque, J. D. Eastman, M. Endl, P. Figueira, E. B. Ford, D. Foreman-Mackey, P. Fournier, G. Fürész, B. S. Gaudi, P. C. Gregory, F. Grundahl, A. P. Hatzes, G. Hébrard, E. Herrero, D. W. Hogg, A. W. Howard, J. A. Johnson, P. Jorden, C. A. Jurgenson, D. W. Latham, G. Laughlin, T. J. Loredo, C. Lovis, S. Mahadevan, T. M. McCracken, F. Pepe, M. Perez, D. F. Phillips, P. P. Plavchan, L. Prato, A. Quirrenbach, A. Reiners, P. Robertson, N. C. Santos, D. Sawyer, D. Segransan, A. Sozzetti, T. Steinmetz, A. Szentgyorgyi, S. Udry, J. A. Valenti, S. X. Wang, R. A. Wittenmyer, and J. T. Wright, *PASP* **128**, 066001 (2016).
 - [5] X. Dumusque, F. Pepe, C. Lovis, and D. W. Latham, *ApJ* **808**, 171 (2015).
 - [6] V. Rajpaul, S. Aigrain, and S. Roberts, *MNRAS* **456**, L6 (2016).
 - [7] X. Dumusque, F. Pepe, C. Lovis, D. Segransan, J. Sahlmann, W. Benz, F. Bouchy, M. Mayor, D. Queloz, N. Santos, and S. Udry, *Nature* **491**, 207 (2012).
 - [8] C. J. Grillmair, D. Charbonneau, A. Burrows, L. Armus, J. Stauffer, V. Meadows, J. V. Cleve, and D. Levine, *ApJL* **658**, L115 (2007).
 - [9] K. O. Todorov, D. Deming, A. Burrows, and C. J. Grillmair, *ApJ* **796**, 100 (2014).
 - [10] S. Millholland, S. Wang, and G. Laughlin, *ApJL* **823**, L7 (2016).
 - [11] S. Agarwal, W. Moon, and J. S. Wettlaufer, *RSPSA* **468**, 2416 (2012).
 - [12] In regular MF-DFA, rather than using temporal-weighting, n th order polynomials $y_\nu(i)$ are used to approximate $Y(i)$ within a fixed window, without reference to points in the profile outside that window.
 - [13] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, *PhyA* **316**, 87 (2002).
 - [14] G. Rangarajan and M. Ding, *PhRvE* **61**, 4991 (2000).
 - [15] F. Bouchy, S. Udry, M. Mayor, C. Moutou, F. Pont, N. Iribarne, R. Da Silva, S. I. ky, D. Queloz, N. C. Santos, D. S. an, and S. Zucker, *A&A* **444**, L15 (2005).
 - [16] X. Dumusque, I. Boisse, and N. C. Santos, *Astrophys. J.* **796**, 132 (2014), [arXiv:1409.3594 \[astro-ph.SR\]](https://arxiv.org/abs/1409.3594).
 - [17] M. Mayor, F. Pepe, D. Queloz, F. Bouchy, G. Rupprecht, G. Lo Curto, G. Avila, W. Benz, J.-L. Bertaux, X. Bonfils, T. Dall, H. Dekker, B. Delabre, W. Eckert, M. Fleury, A. Gilliotte, D. Gojak, J. C. Guzman, D. Kohler, J.-L. Lizon, A. Longinotti, C. Lovis, D. Megevand, L. Pasquini, J. Reyes, J.-P. Sivan, D. Sosnowska, R. Soto, S. Udry, A. van Kesteren, L. Weber, and U. Weilenmann, *Msngr* **114**, 20 (2003).
 - [18] Triaud, A. H. M. J., Queloz, D., Bouchy, F., Moutou, C., Cameron, A. C., Claret, A., Barge, P., Benz, W., Deleuil, M., Guillot, T., Hébrard, G., Lecavelier des Étangs, A., Lovis, C., Mayor, M., Pepe, F., and Udry, S., *A&A* **506**, 377 (2009).

- [19] Wyttenbach, A., Ehrenreich, D., Lovis, C., Udry, S., and Pepe, F., *A&A* **577**, A62 (2015).
- [20] K. Horne, *PASP* **98**, 609 (1986).
- [21] G. Torres, J. N. Winn, and M. J. Holman, *ApJ* **677**, 1324 (2008).
- [22] S. Agarwal and J. S. Wettlaufer, *PhLA* **380**, 142 (2016).
- [23] Spitzer, *Modification to observed pointing wobble during staring observations*, Tech. Rep. (Caltech, 2010).
- [24] A. Collier Cameron, D. M. Wilson, R. G. West, L. Hebb, X. B. Wang, S. Aigrain, F. Bouchy, D. J. Christian, W. I. Clarkson, B. Enoch, M. Esposito, E. Guenther, C. A. Haswell, G. Hébrard, C. Hellier, K. Horne, J. Irwin, S. R. Kane, B. Loeillet, T. A. Lister, P. Maxted, M. Mayor, C. Moutou, N. Parley, D. Pollacco, F. Pont, D. Queloz, R. Ryans, I. Skillen, R. A. Street, S. Udry, and P. J. Wheatley, *MNRAS* **380**, 1230 (2007).
- [25] S. Seager and G. Mallén-Ornelas, *ApJ* **585**, 1038 (2003).
- [26] F. Pont, R. L. Gilliland, C. Moutou, D. Charbonneau, F. Bouchy, T. M. Brown, M. Mayor, D. Queloz, N. Santos, and S. Udry, *A&A* **476**, 1347 (2007).